REVISITING THE FORWARD EQUATIONS FOR INHOMOGENEOUS SEMI-MARKOV PROCESSES

ALEXANDER SOKOL

ABSTRACT. In this paper, we consider a class of inhomogeneous semi-Markov processes directly based on intensity processes for marked point processes. We show that this class satisfies the semi-Markov properties defined elsewhere in the literature. We use the marked point process setting to derive strong upper bounds on various probabilities for semi-Markov processes. Using these bounds, we rigorously prove for the case of countably infinite state space that the transition intensities are right-derivatives of the transition probabilities, and we prove for the case of finite state space that the transition probabilities satisfy the forward equations, requiring only right-continuity of the transition intensities in the time and duration arguments and a boundedness condition. We also show relationships between several classes of semi-Markov processes considered in the literature, and we prove an integral representation for the left derivatives of the transition probabilities in the duration parameter.

1. Introduction

Since the introduction of semi-Markov processes in [18] and [24], this class of stochastic processes have been thoroughly developed and applied in many fields of study. Initially, the semi-Markov processes studied were homogeneous semi-Markov processes, see e.g. [22, 23, 7, 4]. Intuitively, these are similar to homogeneous Markov processes, except that the intensities of jumps depend on the amount of time spent by the process in its current state. This duration dependency is at the center of semi-Markov theory.

Inhomogeneous semi-Markov processes, where the intensity for jumps depend on both time and the time spent by the process in its current state, have been studied as well, see e.g. [10, 14, 25]. Here, the theory becomes more complex, but also allows for more flexible modeling.

Date: April 14, 2015.
2010 *Mathematics Subject Classification.* Primary 60K15; Secondary 60J25, 60G55, 35R09.
Key words and phrases. Semi-Markov process, Strong Markov process, Intensity, PDE.

The strength of semi-Markov theory is its ability to model systems where the change of the state depends on the amount spent in its current state. Many real-world systems exhibit this type of behaviour. One obvious example is in the context of disability insurance, where the intensity for recovery depends on the amount of time spent in a disabled state: Intuitively, if one does not recover within a few years, it is considerably less likely that one will ever recover. Applications of semi-Markov theory to disability insurance, and life insurance in general, can be found in e.g. [3, 10, 9, 2]. Semi-Markov models find applications elsewhere as well, in fields as diverse as for example wind speed modeling, high frequency finance, tourism movements and credit risk, see [6, 5, 17, 19].

The theoretical literature on semi-Markov processes is considerable, covering e.g. [22, 7, 11, 10, 1, 20, 9] and much more, see the books [15, 16] for thorough reference lists. The majority of the literature, however, focuses on the homogeneous case, neglecting the inhomogeneous case. Furthermore, sometimes the interest in applications take precedence to rigor and mathematical precision, with e.g. [10] stating: *We shall give a few proofs, but not where the results can be argued by "direct reasoning" (i. e., by intuition) . . . All this means that we shall sweep some interesting mathematical problems under the rug.*

In this paper, we revisit some fundamental results for semi-Markov processes, namely the definition of semi-Markov processes, related Markov properties, and the characterization of intensities and transition probabilities. In particular, we give precise conditions and a rigorous proof for the transition probabilities to satisfy the forward partial integro-differential equations, utilizing a direct proof based on estimates on the probabilities of multiple jumps in a short time interval. The forward equations is one example of a result related to the citation by [10] above: It is intuitively quite clear that the equations should hold under "sufficient regularity conditions", but obtaining a rigorous proof of this is challenging.

All of our results cover the inhomogeneous case. Furthermore, we also consider the case of countably infinite state spaces when possible. Our results not only infuse known results with new and improved proof methodologies, but also generalize known results. Specifically, our contributions are as follows:

(1). We prove a theorem on the relative sizes of various classes of semi-Markov processes considered in the literature.
(2). We prove an upper bound for the conditional probability of a semi-Markov process making two jumps in a small interval, Lemma 3.1, and demonstrate in our proofs that this lemma can be used as a key tool for rigorous proofs of various analytical properties of semi-Markov processes.
(3). In the case of a countable state space, we give a precise sufficient condition on the intensities of a semi-Markov process for being the derivatives of the transition probabilities.

(4). In the case of a finite state space, we give a precise sufficient condition on the intensities of a semi-Markov process for satisfying the forward partial integro-differential equations.

The remainder of the paper is organized as follows. In Section 2, we review the various notions of semi-Markov processes considered in the literature, we introduce a class of semi-Markov processes with extra regularity, and we prove a theorem on the relative sizes of various classes of semi-Markov processes. In Section 3, we prove a bound on the probability of a semi-Markov process making two jumps in a short interval of time, and we give sufficient conditions for the transition intensities to be the right derivatives of the transition probabilities. In Section 4, we give sufficient conditions for the transition intensities to satisfy the forward equations. Finally, in Section 5, we discuss our results and consider opportunities for further research. Proofs can be found in Appendix A.

2. Weak and strong classes of semi-Markov processes

Several definitions of the notion of a semi-Markov process exist in the literature, varying in their strictness. In this section, we review the definitions used in the literature, and we introduce a class of semi-Markov processes with extra regularity, using intensity processes for marked point processes. Furthermore, we prove a result on the relative sizes of various types of semi-Markov process.

We work in the context of a sequence of random variables $(Y_n, S_n)_{n\geq 0}$, where $S_0 = 0$ and $S_n > 0$ for all $n \geq 1$ and Y_n takes its values in E for all $n \geq 0$, where $(E, \mathcal{E})$ is some measurable space. We assume that $Y_{n+1} \neq Y_n$ for all $n \geq 0$. We then also define $T_n = \sum_{k=0}^{n} S_k$ for $n \geq 0$. In the context of a marked point process, see [12], we think of $(T_n)_{n\geq 1}$ as the event times and of Y_n as the marks, with T_0 being time zero and Y_0 being the initial mark. We then also define $Z_t = Y_n$ for $T_n \leq t < T_{n+1}$, and define $U_t = t - \sup\{0 \leq s \leq t | Z_s \neq Z_t\}$. We refer to U as the duration process of Z. Intuitively, U_t measures the amount of time spent by Z_t in its current state. Here, it is implicit that if $T_{n+1} = \infty$ for some n, then $Z_t = Y_n$ for all $t \geq T_n$. Note that using this definition, it is immediate that Z has cadlag sample paths.

One formulation of the semi-Markov property often seen is given directly through the dependence structure of $(Y_n, S_n)_{n\geq 0}$. In [22], the process Z is defined to be a semi-Markov process when

$$P(Y_n = j, S_n \leq t | Y_0, S_0, Y_1, S_1, \ldots, S_{n-1}, Y_{n-1}) = P_{Y_{n-1},j}(t) \tag{2.1}$$

for some family $P(i, \cdot)$ of distributions on E, where $P_{ij}(t) = P(i, \{j\} \times [0, t])$. Here, then, $P_{ij}(t)$ can be interpreted as the probability of transitioning from i to j, having stayed in the previous state for a duration of less than or equal to t. This definition corresponds to a time-homogeneous case, and essentially states

that (Y_n, S_n) is conditionally independent of $S_0, Y_0, \ldots, S_{n-1}$ given Y_{n-1}. The same definition is also used in [4], where the family $(P(i,\cdot))_{i\in E}$ is referred to a semi-Markovian kernel. The papers [5, 23] also use variants of this definition. Furthermore, [9] defines Z to be a semi-Markov process if $(Y_n, T_n)_{n\geq 0}$ is a Markov process. The majority of authors take the case of finite state space E as their main interest. Some exceptions to this are [7, 24, 27], who consider the case of a countable state space, and [26], who allows a general, possibly uncountable state space.

The other main type of definition considers the process (Z, U). In particular, in [27], Z is said to be a semi-Markov process when (Z, U) is a strong homogeneous Markov process. This type of definition is also applied in [3, 2], allowing for time inhomogeneity. In [10] the author endeavors to retain the discussion on an intuitive level, but generally also argues for a definition of semi-Markov processes of the same type.

The definition based on having (Z, U) be a Markov process has several qualities: It encapsulates the central idea of a semi-Markov process, namely dependence on only the current state and the duration spent in that state, it is concise and it is equally coherent for both finite, countable and uncountable state spaces. As stated in the following definition, we will consider this to be the main defining property of a semi-Markov process. For sake of tractability, we require that (Z, U) in fact be a strong Markov process.

Definition 2.1. Let $(E, \mathcal{E})$ be a measurable space, and let Z be a cadlag process taking its values in E. We say that Z is an *inhomogeneous semi-Markov process* if (Z, U) is an inhomogeneous strong Markov process on $E \times \mathbb{R}_+$.

By itself, however, processes such as those given in Definition 2.1 are not sufficiently regular to allow for a rich mathematical theory. We next introduce a more regular type of semi-Markov processes, based on intensity processes for marked point processes. Our claim is that this class of processes is amenable to discuss various types of regularity and will provide a sound framework for rigorous development of theory.

In the following, we assume given for each $i, j \in E$ with $i \neq j$ a measurable mapping $q_{ij} : \mathbb{R}_+^2 \to \mathbb{R}_+$. We also define $q_i : \mathbb{R}_+^2 \to \mathbb{R}_+$ for $i \in E$ by

$$q_i(t, u) = \sum_{j \neq i} q_{ij}(t, u), \tag{2.2}$$

and we assume throughout that

$$\sup_{(t,u)\in K} \sup_{i\in E} q_i(t, u) < \infty \tag{2.3}$$

for all compact subsets K of $\mathbb{R}_+^2$. This will in particular ensure the absence of explosion for all processes under consideration. The requirement (2.3) may appear

strict at first sight, but really states little more than that all intensities for making jumps are bounded simultaneously on finite intervals of time. In the case where E is finite, the supremum over E is of course of no impact. We furthermore define $q_{ii}(t,u) = -q_i(t,u)$, and we let $Q(t,u)$ denote the $E \times E$ matrix whose entry for the i'th row and j'th column is $q_{ij}(t,u)$.

Definition 2.2. We say that a marked point process process Z with countable state space E is an *inhomogeneous semi-Markov process with intensities* if there exists a family of mappings $q_{ij} : \mathbb{R}_+^2 \to \mathbb{R}_+$ for $i \neq j$ satisfying (2.3) such that Z has intensity process λ given by $\lambda_t(i) = 1_{(Z_{t-} \neq i)} q_{Z_{t-}i}(t, U_{t-})$ for $i \in E$, where we assume that each q_i is Lebesgue integrable on compact subsets of $\mathbb{R}_+^2$.

Note that by Corollary 4.4.4 of [12], the bound (2.3) ensures that explosion does not occur and thus Definition 2.2 is not vacuous. In Theorem 2.3, we show that the processes defined in Definition 2.2 are in fact semi-Markov processes in the sense of Definition 2.1. The requirement that q_i is Lebesgue integrable on compact subsets of $\mathbb{R}_+^2$ is made to ensure that the resulting intensity process $(t,i) \mapsto \lambda_t(i)$ in fact is integrable and thus allows for a corresponding compensator process. The interpretation of Definition 2.2 is straightforward: When the process Z has been in state i for a duration u at time t, the intensity for making a jump to a state $j \neq i$ is $q_{ij}(t,u)$.

We now show a result, Theorem 2.4, about the relative strengths of various notions of semi-Markov processes. The result that Definition 2.2 implies Definition 2.1 is sufficiently important to us to be stated as a separate theorem.

Theorem 2.3. *Assume that Z is an inhomogeneous semi-Markov process with intensities. Then (Z,U) is an inhomogeneous strong Markov process on $E \times \mathbb{R}_+$.*

Theorem 2.4. *Let Z be a cadlag stochastic process with countable state space E, let (T_n) be the jump times of Z, and let U be the duration process of Z. Consider the following four statements:*

(1). *The process Z is an inhomogeneous semi-Markov process with intensities.*
(2). *The process Z is an inhomogeneous semi-Markov process.*
(3). *(Y_n, S_n) is conditionally independent of $S_0, Y_0, \ldots, S_{n-1}$ given Y_{n-1}.*
(4). *$(Y_n, T_n)_{n \geq 0}$ is a discrete-time Markov chain.*

Here, (1) refers to Definition 2.2, and (2) refers to definition Definition 2.1. It then holds that (1) *implies* (2), (2) *implies* (3), *and* (3) *implies* (4).

In the next section, we consider results characterizing the distribution of a semi-Markov process with intensities, in particular characterizing the transition probabilities of the process.

3. Intensities as derivatives of transition probabilities

Our objective in this section will be to rigorously obtain sufficient regularity criteria for the intensity functions q_{ij} to ensure that they can be obtained as the right derivatives of the transition probabilities of the semi-Markov process. Furthermore, we will show that under the same regularity conditions, the convergence of the difference quotients occur under the presence of a dominating bound. This will be essential for our later proof that the transition probabilities satisfy the forward equations.

In the following, let Z be a semi-Markov process with intensities q_{ij} for $i \neq j$. As in the previous section, we assume that the intensities satisfy (2.3). As the process (Z, U) is a Markov process by Theorem 2.3, we may associate to it a family of transition probabilities $P_{s,t}((i,u), \cdot)$ on $E \times \mathbb{R}_+$ such that

$$P((Z_t, U_t) \in A \times B \mid \mathcal{F}_s) = P_{s,t}(Z_s, U_s, A \times B), \tag{3.1}$$

where $A \subseteq E$ and $B \in \mathcal{B}_+$, with $\mathcal{B}_+$ denoting the Borel σ-algebra on $\mathbb{R}_+$. Consistently with the notation outlined in [10] and also used in [2], we define, for $i, j \in E$ and $s, t, u, d \geq 0$ with $s \leq t$,

$$p_{ij}(s,t,u,A) = P_{s,t}(i, u, \{j\} \times A), \tag{3.2}$$

$$p_{ij}(s,t,u,d) = p_{ij}(s,t,u,[0,d]), \tag{3.3}$$

$$p_{ij}(s,t,u) = p_{ij}(s,t,u,\mathbb{R}_+), \tag{3.4}$$

where $A \in \mathcal{B}_+$. We then think of $p_{ij}(s,t,u,A)$ as the probability of transitioning from i to j from time s to time t when the duration at time s is u, and requiring that the duration at time t is in A. We then also similarly think of $p_{ij}(s,t,u)$ as the probability of transitioning from i to j from time s to time t when the duration at time s is u. Finally, as a function of d, $p_{ij}(s,t,u,d)$ is the cumulative mass function for the measure $A \mapsto p_{ij}(s,t,u,A)$.

Our objective is to show that the intensities q_{ij} are right derivatives of transition probabilities. To this end, we first state two lemmas of independent interest to be used in the proof. For use in the following, we define

$$N_t = \sum_{0<s\leq t} 1_{(Z_{s-} \neq Z_s)}, \tag{3.5}$$

the counting process for counting the jumps of Z. Also, when Q is some $E \times E$ matrix, we let $\|Q\|_\infty$ denote the supremum norm of the matrix. Note that even when E is infinite, it always holds for $t, u \geq 0$ that

$$\sup_{(t,u)\in K} \|Q(t,u)\|_\infty = \sup_{(t,u)\in K} \sup_{i,j\in E} |q_{ij}(t,u)| = \sup_{(t,u)\in K} \sup_{i\in E} q_i(t,u) < \infty, \tag{3.6}$$

for compact subsets K of $\mathbb{R}_+^2$, due to the condition (2.3). Finally, we say that the intensities (q_{ij}) are right-continuous if it holds for all $i \neq j$ that $q_{ij}(t+h, u+k)$ tends to $q_{ij}(t,u)$ whenever h and k tends to zero from above.

Lemma 3.1. *It holds that*

$$\begin{aligned} &P(N_{t+h}-N_t \geq 2 | Z_t = i, U_t = u) \\ &\leq \int_t^{t+h} \int_s^{t+h} \|Q(s, u+s-t)\|_\infty \|Q(v, v-s)\|_\infty \,\mathrm{d}v\,\mathrm{d}s. \end{aligned} \tag{3.7}$$

Lemma 3.2. *It holds that*

$$\lim_{h\to 0} \frac{1}{h} P(N_{t+h}-N_t \geq 2 | Z_t = i, U_t = u) = 0. \tag{3.8}$$

Lemma 3.2 essentially shows that for semi-Markov processes satisfying the bound (2.3), the probability of having two jumps in a small time interval tends to zero at a faster than linear rate. This is an important regularity property which is one of the key properties allowing us to prove our results on semi-Markov processes. The lemma follows immediately from Lemma 3.1. Lemma 3.1 is a stronger bound, allowing for a variety of derived bounds on probabilities for semi-Markov processes.

We are now ready to state our two main results of this section. Theorem 3.4 shows that the intensities are limits of transition probabilities, in accordance with our intuitive understanding of intensities. Theorem 3.3 shows that the convergence in Theorem 3.4 occurs under the presence of a dominating bound. This latter result is used in the proof of Theorem 3.4 and will also be essential for our rigorous proof of the forward equations in the next section. In the statement of the theorems, we let I denote the $E \times E$ identity matrix.

Theorem 3.3. *If the intensities are right-continuous, there exists a measurable mapping $C : \mathbb{R}_+^2 \to \mathbb{R}_+$, bounded on compacts, such that there is $\varepsilon > 0$ with the property that for any $t, u \geq 0$ and $0 \leq h \leq \varepsilon$, we have*

$$\sup_{i\in E} \sum_{j\in E} \left| \frac{1}{h}(p_{ij}(t, t+h, u) - I_{ij}) \right| \leq C(t, u). \tag{3.9}$$

Theorem 3.4. *If the intensities are right-continuous, it holds that*

$$\lim_{h\to 0} \frac{1}{h}(P(t, t+h, u) - I) = Q(t, u). \tag{3.10}$$

4. The forward equations

Next, we turn to the characterization of transition probabilities. For homogeneous Markov processes, a closed-form expression for the transition probabilities is available to us through the matrix exponential, see e.g. Section 2.8 of [21], and for inhomogeneous Markov processes, the transition probabilities can be characterized through two different systems of multidimensional ODEs, namely the forward and backward equations, as proven in [8]. In the semi-Markov case, the transition probabilities also satisfy forward and backward systems of equations. In the

case of the forward equations, these take the form of a system of partial integro-differential equations. The existence of these equations for semi-Markov processes is well known. That the transition probabilities satisfy both the forward and backward equations is proved very rigorously in [9] for the case of a homogeneous semi-Markov process. A variant of the forward equations for the inhomogeneous case is stated in [2], where it forms the basis for numerical calculations of cashflows in life insurance.

In this section, we use the results from Section 3 to give a direct and rigorous proof of a sufficient condition on the intensities for the transition probabilities to satisfy the forward equations. We will also show that this set of equations can be rewritten as a set of ordinary integro-differential equations in two different ways.

Throughout this section, we assume that E *is finite and that the transition probabilities are right-continuous.* The condition (2.3) also remains in force, but is now equivalent to the simpler condition

$$\sup_{(t,u)\in K} q_i(t,u) < \infty. \tag{4.1}$$

Before proving that the transition probabilities satisfy the forward equations, we require two lemmas of independent interest. The first lemma states that in an asymptotic sense, no double jumps back and forth are made over a small time period by a semi-Markov chain. Here, we define $p_{ij}(s,t,u,d-) = \lim_{h\to 0^+} p_{ij}(s,t,u,d-h)$ for $d > 0$.

Lemma 4.1. *It holds that*

$$\lim_{h\to 0} \frac{1}{h} p_{ii}(t,t+h,u,(u+h)-) = 0. \tag{4.2}$$

Furthermore, there exists a measurable mapping $C : \mathbb{R}_+^2 \to \mathbb{R}_+$, bounded on compacts, such that there is $\varepsilon > 0$ with the property that for any $t, u \geq 0$ and $0 \leq h \leq \varepsilon$, we have

$$\frac{1}{h} p_{ii}(t,t+h,u,(u+h)-) \leq C(t,u). \tag{4.3}$$

Note that the result in Lemma 4.1 cannot be extended to cover $p_{ii}(t,t+h,u,u+h)$ instead of $p_{ii}(t,t+h,u,(u+h)-)$, since $p_{ii}(t,t+h,u,\cdot)$ is concentrated on $[0,u+h]$ and so it holds that

$$\begin{aligned} \lim_{h\to 0} \frac{1}{h} p_{ii}(t,t+h,u,u+h) &= \lim_{h\to 0} \frac{1}{h} p_{ii}(t,t+h,u) \\ &= \lim_{h\to 0} \frac{1}{h}(p_{ii}(t,t+h,u) - 1) + \frac{1}{h}, \end{aligned}$$

which is infinite, according to Theorem 3.4. Essentially, this shows that in an asymptotic sense, the only mass for $p_{ii}(t,t+h,u,\cdot)$ is a point mass in $u+h$, corresponding to no jumps being made when transitioning from i to i from time t to time $t+h$.

We are now ready to begin work on proving the differential properties of the transition probabilities. We begin by showing that the transition probabilities are differentiable from the left in the final parameter, and we derive an explicit expression for the derivative. With this lemma in hand, we will be able to prove our main result on the forward equations.

Lemma 4.2. *Let $i, j \in E$, let $0 \leq s \leq t$ and let $u \geq 0$. It holds that the mapping $d \mapsto p_{ij}(s,t,u,d)$ is differentiable from the left for $d > 0$. Furthermore, for $i \neq j$ and $d > t-s$, the derivative is zero, for $i \neq j$ and $d \leq s-t$, the derivative is*

$$\frac{\mathrm{d}p_{ij}}{\mathrm{d}d}(s,t,u,d) = \sum_{k \neq j} \int_0^{u+t-d-s} a_{kj}(t,d,u,v) p_{ik}(s, t-d, u, \, \mathrm{d}v). \tag{4.4}$$

where

$$a_{kj}(t,d,u,v) = q_{kj}(t-d,v) \exp\left(-\int_{t-d}^{t} q_j(r, r-(t-d)) \, \mathrm{d}r\right). \tag{4.5}$$

The intuitive explanation for the formula (4.4) is as follows. The derivative is the limit as h tends to zero from above of

$$\begin{aligned} \frac{1}{h}\left(p_{ij}(s,t,u,d) - p_{ij}(s,t,u,d-h)\right) &= \frac{1}{h} p_{ij}(s,t,u,(d-h,d]) \\ &= \frac{1}{h} P(Z_t = j, d-h < U_t \leq d | Z_s = i, U_s = u). \qquad (4.6) \end{aligned}$$

Now, having $d - h < U_t \leq d$ is equivalent to having a jump being made in the time interval $[t-d, t-d+h)$ and having no jumps made in the time interval $[t-d+h, t]$. Conditioning on $Z_{t-d} = k$ for $k \in E$ yields that the the differential quotient (4.6) approximately is the sum over the states k, and for each state, we sum the density of transitioning from k to j immediately after time $t-d$ and then remaining in state j from time $t-d$ to time t with duration zero at time $t-d$. Furthermore, this is weighted with the probability of transitioning from state i to state k from time s to time $t-d$ with duration u at time s. The term corresponding to $k = j$ vanishes, as being in state j at time $t-d$ would indicate an ultimate duration at time t greater than d. In (4.4), the integration with respect to $p_{ik}(s, t-d, u, \, \mathrm{d}v)$ represents conditioning on the state at time $t-d$, the left factor in (4.5) corresponds to the conditional density of making a jump from k to j immediately after time $t-d$, and the right factor in (4.5) corresponds to the probability of remaining in state j from time $t-d$ to time t with duration zero at time $t-d$.

We are now ready to prove that the transition probabilities satisfy the forward partial integro-differential equations.

Theorem 4.3. *Fix $s \geq 0$ and $u \geq 0$. It holds for $t \geq s$ and $d > 0$ that $p_{ij}(s,t,u,d)$ is differentiable in t from the right, and the derivative is*

$$\frac{\partial p_{ij}}{\partial t}(s,t,u,d) = \sum_{k \neq j} \int_0^{u+t-s} q_{kj}(t,v) p_{ik}(s,t,u,\,\mathrm{d}v) + \int_0^d q_{jj}(t,v) p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{\partial p_{ij}}{\partial d}(s,t,u,d), \tag{4.7}$$

where the partial derivative with respect to d is the derivative from the left.

Theorem 4.3 yields the semi-Markovian analogue of the Kolmogorov forward equations for Markov processes. As an immediate corollary of Theorem 4.3, we may also derive a system of ordinary integro-differential equations for the transition probabilities. Applying the chain rule, it holds that

$$\frac{\partial}{\partial t} p_{ij}(s,t,u,d+t-s) = \frac{\partial p_{ij}}{\partial t}(s,t,u,d+t-s) + \frac{\partial p_{ij}}{\partial d}(s,t,u,d+t-s), \tag{4.8}$$

and applying Theorem 4.3 in (4.8), we obtain

$$\frac{\partial}{\partial t} p_{ij}(s,t,u,d+t-s) = \sum_{k \neq j} \int_0^{u+t-s} q_{kj}(t,v) p_{ik}(s,t,u,\,\mathrm{d}v) + \int_0^{d+t-s} q_{jj}(t,v) p_{ij}(s,t,u,\,\mathrm{d}v). \tag{4.9}$$

Note that in (4.8) and (4.9), the term

$$\frac{\partial}{\partial t} p_{ij}(s,t,u,d+t-s) \tag{4.10}$$

refers to differentiation of the composite mapping $p_{ij}(s,t,u,d+t-s)$, and not to the derivative of $p_{ij}(s,t,u,d)$ evaluated in $(s,t,u,d+t-s)$. Also note that in the case of a Markov process, where $q_{ij}(t,v) = q_{ij}(t)$ for some $q_{ij} : \mathbb{R}_+ \to \mathbb{R}_+$, we may let $u = d$ in (4.9) and immediately recover the classical Kolmogorov forward equation from [8].

Furthermore, we may also insert our expression (4.4) for the derivative in d of $p_{ij}(s,t,u,d)$ directly into (4.7) to obtain

$$\frac{\partial p_{ij}}{\partial t}(s,t,u,d) = \sum_{k \neq j} \int_0^{u+t-s} q_{kj}(t,v) p_{ik}(s,t,u,\,\mathrm{d}v) + \int_0^d q_{jj}(t,v) p_{ij}(s,t,u,\,\mathrm{d}v) - \sum_{k \neq j} \int_0^{u+t-d-s} a_{kj}(t,d,u,v) p_{ik}(s,t-d,u,\,\mathrm{d}v). \tag{4.11}$$

for the case $i \neq j$ and $d \leq t-s$, with a similar result holding for the diagonal case.

5. Discussion

In this article, we have considered a class of regular semi-Markov processes defined through intensities for a marked point process. We have proven relationships between different notions of semi-Markov processes, and we have shown how to obtain sufficient conditions on the intensities for having the intensities be the right derivatives of the transition probabilities. Finally, we have given sufficient conditions on the intensities for having the transition probabilities satisfy the forward equations, and we have stated the latter equations in three different variants: in one way as a system of partial integro-differential equations, and in two ways as a system of ordinary integro-differential equations. In the course of this, we have also proved a formula for the derivative of the transition probabilities with respect to the duration d.

Our main purpose and focus has been to develop a framework and tools for rigorous analysis of semi-Markov processes. Our results indicate that analysis of semi-Markov processes centered around conditioning arguments for single jumps, combined with applications of bounds for probabilities of several jumps such as Lemma 3.1 and Lemma 3.2, is a fruitful proof methodology.

We feel that there is ample room for improvement of our results and for continued study of semi-Markov processes. Some subjects for further study of semi-Markov processes which we find particularly interesting are as follows:

(1). Assuming sufficient regularity conditions, the transition probabilities also satisfy a different system of equations, the backward partial differential equations, see e.g. [2, 3], and see [9] for a rigorous proof in the homogeneous case. It is of interest to develop rigorous arguments for sufficient conditions on the intensities ensuring the validity of the backward equations.

(2). In terms of the forward equations, it is of interest to weaken the regularity conditions required for the intensities, and to consider the case of countably infinite state spaces. In the latter scenario, several interchanges of limits and summation would have to be argued for separately.

(3). The forward and backward equations can be used for numerical evaluation of various expressions in semi-Markov models, see e.g. [2]. In terms of numerical studies, it would be of interest to compare the computational efficiency and accuracy of numerical solutions of these equation systems.

We hope that our efforts in this paper will inspire further study of the field of semi-Markov processes.

Appendix A. Proofs

A.1. Proofs for Section 2.

Proof of Theorem 2.3. We may assume without loss of generality that Z has deterministic initial state y_0. We begin by proving that the process (Z, U) is a piecewise deterministic Markov process obtained from a marked point process satisfying the sufficient criteria of Theorem 7.3.1 of [12]. In order to do so, we need to specify both the underlying marked point process as well as the transform functions $\phi_{s,t}$ appearing in that theorem.

Define $G = E \times \mathbb{R}_+$. Letting E be endowed with the discrete topology and letting $\mathbb{R}_+$ be endowed with its usual topology, we endow G with the product topology. With $\mathcal{G}$ denoting the corresponding Borel σ-algebra on G, it holds that $\mathcal{G}$ is countably generated and contains all singletons, as is required by the results of Section 7.3 of [12]. Next, let $\tilde{Z}_t = (Z_t, 0)$ and define $\tilde{T}_n = T_n$ and $\tilde{Y}_n = (Y_n, 0)$. We then obtain that $\tilde{Z}$ is a marked point process with mark space $(G, \mathcal{G})$. Furthermore, for $0 \le s \le t$ and with $\tilde{y} = (y, u)$, define $\phi_{s,t} : G \to G$ by $\phi_{s,t}(\tilde{y}) = (y, t - s)$. Also define $\tilde{T}_{\langle t\rangle} = \sup\{\tilde{T}_n | \tilde{T}_n \le t\}$ and define $\tilde{Y}_{\langle t\rangle} = \tilde{Y}_{\tilde{T}_{\langle t\rangle}}$. We then obtain

$$(Z_t, U_t) = (Y_{\langle t\rangle}, t - T_{\langle t\rangle}) = \phi_{T_{\langle t\rangle},t}(\tilde{Y}_{\langle t\rangle}). \tag{A.1}$$

This shows that (Z, U) can be obtained as a piecewise deterministic process as in Theorem 7.3.1 of [12], with underlying marked point process $\tilde{Z}$. It remains to check the requirements on $\tilde{Z}$ of that theorem. In the notation of [12], the distribution of the marked point process $\tilde{Z}$ satisfies

$$\bar{P}^{(n)}_{\tilde{z}_n}(t) = \exp\left(-\int_{t_n}^{t} q_{y_n}(s, s - t_n)\,\mathrm{d}s\right), \tag{A.2}$$

$$\pi^{(n)}_{\tilde{z}_n,t}(\{(j, 0)\}) = \frac{q_{y_n j}(t, t - t_n)}{q_{y_n}(t, t - t_n)}, \tag{A.3}$$

for $t \ge 0$, $n \ge 1$ and $j \in E$ with $(j, 0) \neq \tilde{y}_n$ and with (A.3) the latter being defined to be zero when the denominator is zero, the value of $\pi^{(n)}_{\tilde{z}_n,t}(\{(j, 0)\})$ in this case is irrelevant for the distribution of the marked point process. Here, $\tilde{z}_n = (t_1, \tilde{y}_1, \ldots, t_n, \tilde{y}_n)$ and $\tilde{y}_n = (y_n, 0)$. Also, similar expressions are obtained for the distribution of $(\tilde{T}_1, \tilde{Y}_1)$. With $\tilde{y} = (y, u) \in G$, now define $\tilde{q}_t(\tilde{y}) = q_y(t, u)$. Noting that $\phi_{0,t}(y, u) = (y, t)$, we obtain

$$q_{y_0}(s, s) = \tilde{q}_s(y_0, s) = \tilde{q}_s(\phi_{0,s}(\tilde{y}_0)), \tag{A.4}$$

$$q_{y_n}(s, s - t_n) = \tilde{q}_s(y_n, s - t_n) = \tilde{q}_s(\phi_{t_n,s}(\tilde{y}_n)), \tag{A.5}$$

so that with $t \ge 0$, $(y, u) \in G$ and $C \in \mathcal{G}$ and

$$r_t((y, u), C) = \sum_{(j,v)\in C} 1_{(j\neq y, v=0)} \frac{q_{yj}(t, u)}{q_y(t, u)}, \tag{A.6}$$

we obtain

$$\bar{P}^{(n)}_{\tilde{z}_n}(t) = \exp\left(-\int_{t_n}^{t} \tilde{q}_s(\phi_{t_n,s}(\tilde{y}_n))\,\mathrm{d}s\right), \tag{A.7}$$

$$\pi^{(n)}_{z_n,t}(C) = r_t(\phi_{t_n,t}(\tilde{y}_n), C). \tag{A.8}$$

Finally, we have that

$$\int_t^{t+h} \tilde{q}_s(\phi_{s,t}(\tilde{y}))\,\mathrm{d}s = \int_t^{t+h} q_y(s, t-s)\,\mathrm{d}s \tag{A.9}$$

is finite for small h by assumption, so that (i) of Theorem 7.3.1 of [12] is satisfied. As the remaining requirements (ii) and (iii) are trivially satisfied, we may invoke the theorem and obtain that (Z, U) is a piecewise deterministic Markov process obtained from a marked point process satisfying the sufficient criteria of Theorem 7.3.1 of [12]. Theorem 7.5.1 of [12] then furthermore shows that (Z, U) is a strong inhomogeneous Markov process. □

Lemma A.1. *For all $n \geq 1$, it holds that when $T_n < \infty$, we have $Z_{T_n} = Y_n$ and $Z_{T_n-} = Z_{T_{n-1}}$. Furthermore, for $T_n \leq t < T_{n+1}$, it holds that $U_t = t - T_n$. In particular, $U_{T_n-} = S_n$, with $S_n = T_n - T_{n-1}$.*

Proof of Lemma A.1. The result on the values of Z_{T_n} and $Z_{T_{n-1}}$ follow immediately from the pathwise definition of Z in terms of T_n and Y_n. As regards the claims about U, let $T_n \leq t < T_{n+1}$. It then holds that Z is constant on $[T_n, t]$, while it holds that $Z_{T_n-} = Y_{n-1} \neq Y_n = Z_{T_n}$, so Z changes its value at T_n. Therefore, we have

$$\sup\{0 \leq s \leq t | Z_s \neq Z_t\} = T_n, \tag{A.10}$$

and so $U_{T_n} = t - T_n$. In particular, it follows that

$$U_{T_n-} = \lim_{h\to 0^+} U_{T_n-h} = \lim_{h\to 0^+} (T_n - h) - T_{n-1} = T_n - T_{n-1} = S_n, \tag{A.11}$$

as required. □

Proof of Theorem 2.4. That (1) implies (2) is the content of Theorem 2.3. To prove that (2) implies (3), assume that (Z, U) is an inhomogeneous Markov process. Note that as $S_0 = 0$, it suffices to prove that (Y_n, S_n) is conditionally independent of $Y_0, S_1 \ldots, Y_{n-2}, S_{n-1}$ given Y_{n-1}. Noting that we always have $U_{T_{n-1}} = 0$, we may apply Lemma A.1 and the strong Markov property at the stopping time T_{n-1}, obtaining

$$\begin{aligned}
&P(Y_n = j, S_n \leq t | Y_0 = i_0, S_1 = s_1, \ldots, S_{n-1} = s_{n-1}, Y_{n-1} = i_{n-1}) \\
&= P(Z_{T_n} = j, U_{T_n-} \leq t | Y_0 = i_0, S_1 = s_1, \ldots, S_{n-1} = s_{n-1}, Y_{n-1} = i_{n-1}) \\
&= P(Z_{T_n} = j, U_{T_n-} \leq t | U_{T_{n-1}} = 0, Z_{T_{n-1}} = i_{n-1}) \\
&= P(Y_n = j, S_n \leq t | Y_{n-1} = i_{n-1}),
\end{aligned} \tag{A.12}$$

which shows the desired conditional independence statement. As regards the proof that (3) implies (4), we note that with $s_k = t_k - t_{k-1}$ for $1 \leq k \leq n-1$, it holds

with $a = t - t_{n-1}$ that

$$\begin{aligned} &P(Y_n = j, T_n \le t | Y_0 = i_0, T_1 = t_1, \dots, T_{n-1} = t_{n-1}, Y_{n-1} = i_{n-1}) \\ &= P(Y_n = j, S_n \le a | Y_0 = i_0, S_1 = s_1, \dots, S_{n-1} = s_{n-1}, Y_{n-1} = i_{n-1}) \\ &= P(Y_n = j, S_n \le t - t_{n-1} | Y_{n-1} = i_{n-1}). \end{aligned} \tag{A.13}$$

As this does not depend on $t_1, \dots, t_{n-2}$, and neither on $i_0, \dots, i_{n-2}$, we conclude that (Y_n, T_n) is conditionally independent of $Y_0, T_0, \dots, Y_{n-2}, T_{n-2}$ given Y_{n-1} and T_{n-1}, yielding the desired discrete-time Markov property. □

A.2. Proofs for Section 3.

Before proving the main results of Section 3, we first prove a series of lemmas. To this end, we fix some notation and recall some facts about hazard functions and hazard measures. First, for $t \ge 0$ and any stopping time S with respect to the filtration $(\mathcal{F}_t)$ induced by Z, let $T(S)$ denote the first jump strictly after time S. Note that as we have $T(S) = \inf\{T_n | T_n > S\} = \inf (T_n)_{(T_n > S)}$ and $(T_n > S) \in \mathcal{F}_{T_n}$, see e.g. Section I.1b of [13], it holds that $T(S)$ is a stopping time for any S. Also, recall that for a distribution μ on $\mathbb{R}_+ \cup \{\infty\}$ whose restriction to $\mathbb{R}_+$ has density f with respect to the Lebesgue measure, we may define the hazard function $h : \mathbb{R}_+ \to \mathbb{R}_+$ of μ by $h(t) = f(t)/\mu((t, \infty])$ whenever the denominator is nonzero, otherwise we let $h(t) = 0$, and it then holds that

$$\mu((t, \infty]) = \exp\left(-\int_0^t h(s)\,\mathrm{d}s\right), \tag{A.14}$$

$$f(t) = h(t) \exp\left(-\int_0^t h(s)\,\mathrm{d}s\right). \tag{A.15}$$

For distributions which are not absolutely continuous, the more general construct of a hazard measure can be applied, see Section 4.1 of [12]. The following four lemmas yield distributions and hazard functions for various variables related to the distribution of Z. In the following, we let $D(E)$ denote the set of cadlag mappings from $\mathbb{R}_+ \to E$, and generally apply the notation of [12] when considering distributional properties of marked point processes.

Lemma A.2. *The conditional distribution of T_{n+1} given $T_0, Y_0, \dots, T_n, Y_n$ has hazard function $h : \mathbb{R}_+ \to \mathbb{R}_+$ given by $h(t) = 1_{(t > T_n)} q_{Y_n}(t, t - T_n)$.*

Proof of A.2. With $z_n = (t_0, y_0, \dots, t_n, y_n)$, let $P^{(n)}_{z_n}$ denote the conditional distribution of T_{n+1} given $(T_0, Y_0, \dots, T_n, Y_n) = z_n$, and let $\Lambda : D(E) \to D(E)$ denote the total compensator of Z. From the explicit representation of the compensator of counting processes given in Section 4.3 of [12], we obtain that the hazard measure

of the conditional distribution is concentrated on (t_n, ∞) and given by

$$\nu_{z_n}^{(n)}((t_n, t]) = \Lambda_t(z) - \Lambda_{t_n}(z) = \sum_{j \neq y_n} \int_{t_n}^{t} q_{y_n j}(s, u_{s-}) \, \mathrm{d}s$$
$$= \int_{t_n}^{t} q_{y_n}(s, s - t_n) \, \mathrm{d}s, \tag{A.16}$$

see (4.47) of [12], for paths z such that only n jumps are made on $[0, t)$ and corresponding to having jump times $t_1, \ldots, t_n$ with destination states $y_1, \ldots, y_n$. As the hazard measure is absolutely continuous, we obtain that the hazard function for the distribution exists and is given by $t \mapsto 1_{(t > t_n)} q_{y_n}(t, t - t_n)$, the Radon-Nikodym derivative of the hazard measure. This proves the result. ☐

Lemma A.3. *The conditional distribution of Y_{n+1} given $T_0, Y_0, \ldots, Y_n, T_{n+1}$ is almost surely given by*

$$P(Y_n = j | T_0, Y_0, \ldots, Y_n, T_{n+1}) = \frac{q_{Y_n j}(t, t - T_{n+1})}{q_{Y_n}(t, t - T_{n+1})}, \tag{A.17}$$

understanding that on the almost sure set when the above holds, the denominator is nonzero.

Proof of A.3. With $\mathcal{P}(E)$ denoting the power set of E, let $\pi_{z_n, t}^{(n)} : \mathcal{P}(E) \to [0, 1]$ denote the conditional distribution of Y_{n+1} given $(T_0, Y_0, \ldots, T_n, Y_n, T_{n+1}) = (z_n, t)$. Let $\Lambda : D(E) \to D(E)$ denote the total compensator of Z, and let the compensator for the jumps to state i be denoted by $\Lambda^i : D(E) \to D(E)$. By (4.48) of [12], we have that whenever $q_{y_n}(t, t - t_n) \neq 0$, it holds for all $A \subseteq E$ that

$$\pi_{z_n, t}^{(n)}(A) = \sum_{i \in A} \frac{\mathrm{d}\Lambda^i}{\mathrm{d}\Lambda}(t) = \sum_{i \in A} \frac{1_{(y_n \neq i)} q_{y_n i}(t, t - t_n)}{q_{y_n}(t, t - t_n)}, \tag{A.18}$$

which yields the result. ☐

Lemma A.4. *The conditional distribution of T_{n+1} given $Z_t = i$, $U_t = u$ and $N_t = n$ has hazard function $h : \mathbb{R}_+ \to \mathbb{R}_+$ given by $h(s) = 1_{(s > t)} q_i(t, u + s - t)$.*

Proof of Lemma A.4. Let $s \geq 0$. We find that

$$P(T_{n+1} \leq s | Z_t = i, U_t = u, N_t = n)$$
$$= P(T_{n+1} \leq s | Y_n = i, T_n = t - u, T_{n+1} > t)$$
$$= \frac{P(T_{n+1} \leq s, T_{n+1} > t | Y_n = i, T_n = t - u)}{P(T_{n+1} > t | Y_n = i, T_n = t - u)}. \tag{A.19}$$

Now, by Lemma A.2, the distribution of T_{n+1} given $Y_n = i$ and $T_n = t - u$ has hazard function $v \mapsto 1_{(v > t - u)} q_i(v, v - (t - u))$. Therefore, we in particular obtain

$$P(T_{n+1} > t | Y_n = i, T_n = t - u) = \exp\left(\int_{t-u}^{t} q_i(v, v - (t - u)) \, \mathrm{d}v\right), \tag{A.20}$$

and, for $s > t$,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}s} P(T_{n+1} \le s, T_{n+1} > t | Y_n = i, T_n = t-u) \\
&= q_i(s, s-(t-u)) \exp\left(\int_{t-u}^{s} q_i(v, v-(t-u))\,\mathrm{d}v \right). \qquad \text{(A.21)}
\end{aligned}$$

Combining (A.19), (A.20) and (A.21) yields that for $s > t$, it holds that

$$\begin{aligned}
&P(T_{n+1} \le s | Z_t = i, U_t = u, N_t = n) \\
&= q_i(s, s-(t-u)) \exp\left(\int_{t}^{s} q_i(v, v-(t-u))\,\mathrm{d}v \right), \qquad \text{(A.22)}
\end{aligned}$$

proving that the conditional distribution of T_{n+1} given $Z_t = i$, $U_t = u$ and $N_t = n$ has hazard $s \mapsto q_i(t, u+s-t)$ for $s > t$. As the support of the distribution is contained in $[t, \infty)$, this proves the lemma. □

Lemma A.5. *The conditional distribution of $T(t)$ given $Z_t = i$ and $U_t = u$ has hazard function $h : \mathbb{R}_+ \to \mathbb{R}_+$ given by $h(s) = 1_{(s>t)} q_i(t, u+s-t)$.*

Proof of Lemma A.5. Define $g : \mathbb{N}_0 \to [0,1]$ by $g(n) = P(N_t = n | Z_t = i, U_t = u)$. Applying Lemma A.4, it then holds for $s > t$ that

$$\begin{aligned}
P(T(t) \le s | Z_t = i, U_t = u) &= \sum_{n=0}^{\infty} P(T(t) > s | Z_t = i, U_t = u, N_t = n) g(n) \\
&= \sum_{n=0}^{\infty} P(T_{n+1} > s | Z_t = i, U_t = u, N_t = n) g(n) \\
&= \sum_{n=0}^{\infty} \exp\left(\int_t^s q_i(v, u+v-t)\,\mathrm{d}v \right) g(n) \\
&= \exp\left(\int_t^s q_i(v, u+v-t)\,\mathrm{d}v \right), \qquad \text{(A.23)}
\end{aligned}$$

yielding the result. □

From a heuristic perspective, Lemma A.5 allows us some insight into the distribution of Z. The lemma states that given the values of (Z_t, U_t), the distribution next jump has a fixed hazard, independently of the number of jumps made. We are now ready to prove the results stated in Section 3.

Proof of Lemma 3.1. For $n \geq 0$, define $g(n) = P(N_t = n, Z_t = i, U_t = u)$. We then have

$$\begin{aligned} &P(N_{t+h} - N_t \geq 2 | Z_t = i, U_t = u) \\ &= \sum_{n=0}^{\infty} P(N_{t+h} - N_t \geq 2 | Z_t = i, U_t = u, N_t = n) g(n) \\ &= \sum_{n=0}^{\infty} P(t < T_{n+1} < T_{n+2} \leq t + h | Z_t = i, U_t = u, N_t = n) g(n). \end{aligned} \tag{A.24}$$

Now, by Lemma A.4, the conditional distribution of T_{n+1} given $Z_t = i$, $U_t = u$ and $N_t = n$ has hazard function $q_i(s, u + s - t)$ for $s \geq t$. Applying Lemma A.2, we then obtain that the conditional distribution of $Z_{T_{n+1}}$ and T_{n+1} given $Z_t = i$, $U_t = u$ and $N_t = n$ has density

$$f(k, s) = q_{ik}(s, u + s - t) \exp\left(- \int_t^s q_i(v, u + v - t) \, \mathrm{d}v \right). \tag{A.25}$$

for $s \geq t$ and zero otherwise. Next, noting that we always have $U_{T_{n+1}} = 0$, we may use the strong Markov property of (Z, U) at T_{n+1} and obtain

$$\begin{aligned} &P(t < T_{n+1} < T_{n+2} \leq t + h | Z_t = i, U_t = u, N_t = n) \\ &= \sum_{k \in E} \int_t^{t+h} P(s < T_{n+2} \leq t + h | Z_{T_{n+1}} = k, T_{n+1} = s) f(k, s) \, \mathrm{d}s. \end{aligned} \tag{A.26}$$

Here, the distribution of T_{n+2} given $Z_{T_{n+1}} = k$ and $T_{n+1} = s$ has hazard $q_k(v, v-s)$ for $v \geq s$ and zero otherwise, which yields the conditional density

$$f(v | k, s) = q_k(v, v - s) \exp\left(- \int_s^v q_k(r, r - s) \, \mathrm{d}r \right). \tag{A.27}$$

for $v \geq s$ and zero otherwise. We thus obtain

$$\begin{aligned} P(s < T_{n+2} \leq t + h | Z_{T_{n+1}} = k, T_{n+1} = s) &= \int_s^{t+h} f(v | k, s) \, \mathrm{d}v \\ &\leq \int_s^{t+h} \| Q(v, v - s) \|_\infty \, \mathrm{d}v. \end{aligned} \tag{A.28}$$

Inserting this in (A.26), we obtain the bound

$$\begin{aligned} &P(t < T_{n+1} < T_{n+2} \leq t + h | Z_t = i, U_t = u, N_t = n) \\ &\leq \sum_{k \in E} \int_t^{t+h} \int_s^{t+h} \| Q(v, v - s) \|_\infty \| Q(s, u + s - t) \|_\infty \, \mathrm{d}v \, \mathrm{d}s. \end{aligned} \tag{A.29}$$

As this bound is independent of n, we may use it in (A.24) and obtain the required result. $\square$

Proof of Lemma 3.2. By Lemma 3.1, we have

$$\begin{aligned}&\frac{1}{h}P(N_{t+h}-N_t\geq 2|Z_t=i,U_t=u)\\ &\leq h\sup_{t\leq s,v\leq t+h}\|Q(s,u+s-t)\|_\infty\|Q(v,v-s)\|_\infty.\end{aligned}\tag{A.30}$$

Now, by the bound (3.6), we find that $(s,v)\mapsto\|Q(s,v)\|_\infty$ is bounded in both a neighborhood of (t,u) and $(t,0)$. Therefore, the above tends to zero as h tends to zero. □

Proof of Theorem 3.3. We will show the result with $\varepsilon=1$. To obtain (3.9) for this case, we first fix $t,u\geq 0$ and $i\in E$ and seek to bound the sum of $\frac{1}{h}(p_{ij}(t,t+h,u)-I_{ij})$ over $j\neq i$ for $h\leq 1$. Fix $i\neq j$, we then have

$$\begin{aligned}p_{ij}(t,t+h,u)&\leq P(N_{t+h}-N_t=1,Z_{t+h}=j|Z_t=i,U_t=u)\\ &\quad+P(N_{t+h}-N_t\geq 2,Z_{t+h}=j|Z_t=i,U_t=u).\end{aligned}\tag{A.31}$$

Next, note that by Lemma A.3 and Lemma A.5, we have

$$\begin{aligned}&P(N_{t+h}-N_t=1,Z_{t+h}=j|Z_t=i,U_t=u)\\ &\leq\int_t^{t+h}q_{ij}(s,u+s-t)\exp\left(-\int_t^s q_i(v,u+v-t)\,\mathrm{d}v\right)\mathrm{d}s\\ &\leq\int_t^{t+h}q_{ij}(s,u+s-t)\,\mathrm{d}s,\end{aligned}\tag{A.32}$$

since $N_{t+h}-N_t=1$ and $Z_{t+h}=j$ is equivalent to having the next jump strictly after t occurring in the interval $(t,t+h]$ with destination state j, and the jump after that occurring strictly after $t+h$. Applying this bound in (A.31), the monotone convergence theorem yields

$$\begin{aligned}\sum_{j\neq i}p_{ij}(t,t+h,u)&\leq\int_t^{t+h}q_i(s,u+s-t)\,\mathrm{d}s\\ &\quad+P(N_{t+h}-N_t\geq 2|Z_t=i,U_t=u).\end{aligned}\tag{A.33}$$

Here, by Lemma 3.1, we have

$$\begin{aligned}&P(N_{t+h}-N_t\geq 2|Z_t=i,U_t=u)\\ &\leq\int_t^{t+h}\int_s^{t+h}\|Q(s,u+s-t)\|_\infty\|Q(v,v-s)\|_\infty\,\mathrm{d}v\,\mathrm{d}s.\end{aligned}\tag{A.34}$$

Combining (A.32) and (A.34) with (A.31), we obtain for $h\leq 1$ that

$$\begin{aligned}&\sum_{j\neq i}\frac{1}{h}p_{ij}(t,t+h,u)\\ &\leq\frac{1}{h}\int_t^{t+h}\|Q(s,u+s-t)\|_\infty\left(1+\int_s^{t+h}\|Q(v,v-s)\|_\infty\,\mathrm{d}v\right)\mathrm{d}s\\ &\leq\sup_{t\leq s\leq t+1}\|Q(s,u+s-t)\|_\infty\left(1+\sup_{s\leq v\leq t+1}\|Q(v,v-s)\|_\infty\right)\mathrm{d}s.\end{aligned}\tag{A.35}$$

As we also have

$$\left|\frac{1}{h}p_{ii}(t,t+h,u)-1\right| = \sum_{j\neq i}\frac{1}{h}p_{ij}(t,t+h,u), \tag{A.36}$$

we finally obtain

$$\begin{aligned}&\sup_{i\in E}\sum_{j\in E}\left|\frac{1}{h}(p_{ij}(t,t+h,u)-I_{ij})\right| \\ &\leq 2\sup_{t\leq s\leq t+1}\|Q(s,u+s-t)\|_\infty\left(1+\sup_{s\leq v\leq t+1}\|Q(v,v-s)\|_\infty\right)\mathrm{d}s.\end{aligned} \tag{A.37}$$

Letting $C(t,u)$ be the right-hand side of (A.35), we obtain (3.9) for the case $i\neq j$. By the right-continuity of the intensities, this definition of the mapping C is measurable, and by (3.6), it is bounded on compacts. □

Proof of Theorem 3.4. Consider first the case $i\neq j$. Here, we need to show that

$$q_{ij}(t,u) = \lim_{h\to 0}\frac{1}{h}p_{ij}(t,t+h,u), \tag{A.38}$$

To this end, we first make the decomposition

$$\begin{aligned}p_{ij}(t,t+h,u) &= P(N_{t+h}-N_t=1, Z_{t+h}=j|Z_t=i,U_t=u) \\ &\quad + P(N_{t+h}-N_t\geq 2, Z_{t+h}=j|Z_t=i,U_t=u).\end{aligned} \tag{A.39}$$

Here, Lemma 3.2 shows that the second term tends to zero, so it suffices to show that the first term tends to $q_{ij}(t,u)$. To do so, we note that having $N_{t+h}-N_t=1$ and $Z_{t+h}=j$ is equivalent to having $T(t)\leq t+h$, $T(T(t))>t+h$ and $Z_{T(t)}=j$. Defining

$$f(t,u,s) = q_{ij}(s,u+s-t)\exp\left(-\int_t^s q_i(v,u+v-t)\,\mathrm{d}v\right), \tag{A.40}$$

$$a(t,r) = \exp\left(-\int_t^r q_j(v,v)\,\mathrm{d}v\right), \tag{A.41}$$

for $s\geq t$ and $r\geq t$, we find by Lemma A.5 and Lemma A.3 that $s\mapsto f(t,u,s)$ is the conditional density given $Z_t=i$ and $U_t=u$ of making the next jump at time t to state j, and $r\mapsto a(t,r)$ is the survival function for the next jump given $Z_t=j$ and $U_t=0$. Therefore, we obtain

$$\begin{aligned}&= P(N_{t+h}-N_t=1, Z_{t+h}=j|Z_t=i,U_t=u) \\ &= \int_t^{t+h} f(t,u,s)a(s,t+h)\,\mathrm{d}s.\end{aligned} \tag{A.42}$$

As the intensities are assumed to be right-continuous, this yields

$$\lim_{h\to 0}\frac{1}{h}P(N_{t+h}-N_t=1, Z_{t+h}=j|Z_t=i,U_t=u) = q_{ij}(t,u), \tag{A.43}$$

which, combined with (A.39), yields (A.38). It remains to prove for any $i \in E$ that

$$q_{ii}(t,u) = \lim_{h\to 0} \frac{1}{h}(p_{ii}(t,t+h,u)-1). \tag{A.44}$$

In order to obtain this, we simply apply the dominated convergence theorem with the bound $C(t,u)$ from Theorem 3.3 to obtain

$$\begin{aligned}
\lim_{h\to 0} \frac{1}{h}(p_{ii}(t,t+h,u)-1) &= -\lim_{h\to 0} \sum_{j\neq i} \frac{1}{h} p_{ij}(t,t+h,u) \\
&= -\sum_{j\neq i} \lim_{h\to 0} \frac{1}{h} p_{ij}(t,t+h,u) \\
&= -\sum_{j\neq i} q_{ij}(t,u) = q_{ii}(t,u), \qquad \text{(A.45)}
\end{aligned}$$

as required. □

A.3. Proofs for Section 4. We begin by proving a few auxiliary lemmas, including a version of the Chapman-Kolmogorov equations for semi-Markov processes.

Lemma A.6. *It holds for all* $0 \le s \le t$ *that* $U_t - U_s \le t - s$.

Proof of Lemma A.6. Fix $0 \le s \le t$. If there are no event times in $(s,t]$, then Lemma A.1 shows that $U_t - U_s = t - s$. Considering the case where there is an event time in $(s,t]$, let T be the largest such event time, $T = \sup\{T_n | s < T_n \le t\}$. We then obtain

$$U_t - U_s = t - T - U_s \le t - T \le t - s, \tag{A.46}$$

as required, since U is nonnegative. □

Lemma A.7. *Let* $0 \le s \le t$*, and let* $u \ge 0$*. For* $i \neq j$*, it holds that the measure* $p_{ij}(s,t,u,\cdot)$ *is concentrated on the set* $[0,t-s]$*. For* $i \in E$*, it holds that the measure* $p_{ii}(s,t,u,\cdot)$ *is concentrated on* $[0,t-s] \cup \{u, u+t-s\}$.

Proof of Lemma A.7. Consider first the case where $i \neq j$. For $s = t$, it is immediate that $p_{ij}(s,t,u,\cdot)$ is the zero measure, so the result is immediate in this case. Consider instead the case where $s < t$. Here, we have

$$\begin{aligned}
p_{ij}(s,t,u,A) &= P(Z_t = j, U_t \in A | Z_s = i, U_s = u) \\
&= P(Z_s = i, Z_t = j, U_t \in A | Z_s = i, U_s = u). \qquad \text{(A.47)}
\end{aligned}$$

Now, when $Z_s = i$ and $Z_t = j$, with $s < t$, it must hold that there is some event time for Z in the interval $(s, t]$. Therefore, we may apply Lemma A.6 and obtain

$$\begin{aligned}
(Z_s = i, Z_t = j) &\subseteq \cup_{n=1}^{\infty}(s < T_n \leq t, Z_s = i, Z_t = j) \\
&\subseteq \cup_{n=1}^{\infty}(s < T_n \leq t, U_{T_n} = 0, Z_s = i, Z_t = j) \\
&\subseteq \cup_{n=1}^{\infty}(s < T_n \leq t, U_t \leq t - T_n, U_{T_n} = 0, Z_s = i, Z_t = j) \\
&\subseteq (U_t \leq t - s). \qquad \text{(A.48)}
\end{aligned}$$

Using this in (A.47), we obtain

$$\begin{aligned}
p_{ij}(s, t, u, A) &= P(Z_t = j, U_t \in A | Z_s = i, U_s = u) \\
&= P(Z_s = i, Z_t = j, U_t \in A \cap [0, t - s] | Z_s = i, U_s = u), \qquad \text{(A.49)}
\end{aligned}$$

as required. Next, consider a single $i \in E$. In the case where $s = t$, it is immediate that $p_{ii}(s, t, u, \cdot)$ is concentrated on $\{u\}$, and so the result follows. Consider instead $s < t$. In this case, it either holds that there are zero event times in $(s, t]$, or there is an event time. In the latter case, we can use a calculation as in (A.48) to obtain that the duration at time t must be in $[0, u + t - s]$. In the former case, we have

$$\begin{aligned}
&\cap_{n=1}^{\infty} (T_n \notin (s, t], Z_s = i, U_s = u, Z_t = j) \\
&= \cap_{n=1}^{\infty}(T_n \notin (s, t], Z_s = i, U_s = u, Z_t = j, U_t = u + t - s). \qquad \text{(A.50)}
\end{aligned}$$

From this, the result follows. □

The analogue of the Chapman-Kolmogorov equations for semi-Markov processes is encapsulated in the following lemma. Note that the integration domain in (A.51) is assumed to be $[0, u+t-s]$ and not $[0, u+t-s)$. This is important as $p_{ij}(s, r, u, \cdot)$ in the case $i = j$ can have a point mass in $u + t - s$, corresponding to the case where no jumps are made in the interval $(s, r]$.

Lemma A.8. *Fix $0 \leq s \leq r \leq t$. It holds that*

$$p_{ij}(s, t, u, A) = \sum_{k \in E} \int_0^{u+r-s} p_{kj}(r, t, v, A) p_{ik}(s, r, u, \mathrm{d}v). \qquad \text{(A.51)}$$

In particular, we have for $d \geq 0$ that

$$p_{ij}(s, t, u, d) = \sum_{k \in E} \int_0^{u+r-s} p_{kj}(r, t, v, d) p_{ik}(s, r, u, \mathrm{d}v). \qquad \text{(A.52)}$$

Proof of Lemma A.8. By the Chapman-Kolmogorov equations for the inhomogeneous Markov chain (Z, U), we have

$$\begin{aligned}
p_{ij}(s, t, u, A) = P_{s,t}(i, u, \{j\} \times A) &= \int_{E \times \mathbb{R}_+} P_{r,t}(k, v, \{j\} \times A) \,\mathrm{d}P_{s,r}(i, u, \mathrm{d}k, \mathrm{d}v) \\
&= \sum_{k \in E} \int_0^{\infty} P_{r,t}(k, v, \{j\} \times A) \,\mathrm{d}P_{s,r}(i, u, k, \mathrm{d}v) \\
&= \sum_{k \in E} \int_0^{\infty} p_{kj}(r, t, v, A) p_{ik}(s, r, u, \mathrm{d}v). \qquad \text{(A.53)}
\end{aligned}$$

As Lemma A.7 shows that the support of $p_{ik}(s,r,u,\cdot)$ is included in $[0,u+r-s]$, we obtain the result. The identity (A.52) follows from (A.51) by definition. $\square$

Proof of Lemma 4.1. We first note that for Z to remain in state i at time $t+h$ conditionally on being in state i at time t can only happen by changing state an even number of times. In particular, this yields

$$\begin{aligned} p_{ii}(t,t+h,u,(u+h)-) &= P(Z_{t+h}=i, U_{t+h} < u+h | Z_t = i, U_t = u) \\ &\le P(N_{t+h}-N_t = 0, U_{t+h} < u+h | Z_t = i, U_t = u) \\ &\quad + P(N_{t+h}-N_t \ge 2 | Z_t = i, U_t = u). \end{aligned} \tag{A.54}$$

Here, we have

$$\begin{aligned} &P(N_{t+h}-N_t = 0, U_{t+h} < u+h | Z_t = i, U_t = u) \\ &= P(N_{t+h}-N_t = 0, U_t = u, U_{t+h} < u+h | Z_t = i, U_t = u) = 0, \end{aligned} \tag{A.55}$$

since, when making no jumps in the time interval $(t,t+h]$, $U_t = u$ implies that $U_{t+h} = u+h$. As a consequence, we obtain

$$\frac{1}{h} p_{ii}(t,t+h,u,(u+h)-) \le \frac{1}{h} P(N_{t+h}-N_t \ge 2 | Z_t = i, U_t = u). \tag{A.56}$$

The limit statement (4.2) then follows from Lemma 3.2, and the existence of the bound (4.3) follows as in the proof of Theorem 3.3. $\square$

In order to prove Lemma 4.2, we first define

$$r_{ij}(s,u,t) = q_{ij}(t,u+t-s) \exp\left(-\int_s^t q_i(v,u+v-s)\,\mathrm{d}v\right), \tag{A.57}$$

$$R_i(s,u,t) = \exp\left(-\int_s^t q_i(v,u+v-s)\,\mathrm{d}v\right). \tag{A.58}$$

Note that $(j,t) \mapsto r_{ij}(s,u,t)$ is then the conditional density of the next jump and its destination state given $Z_s = i$ and $U_s = u$, and $t \mapsto R_i(s,u,t)$ is the survival function for the next jump given $Z_s = i$ and $U_s = u$.

Proof of Lemma 4.2. First consider the case where $i \neq j$. Note that as $p_{ij}(s,t,u,\cdot)$ is concentrated on $[0,t-s]$ by Lemma A.7, we have that whenever $d-h \ge t-s$, it holds that

$$p_{ij}(s,t,u,d) - p_{ij}(s,t,u,d-h) = p(s,t,u,(d-h,d]) = 0. \tag{A.59}$$

In particular, if $d > t-s$, it is immediate that the derivative from the left exists and is zero. Next, consider the case where $d \le t-s$. As $s \le t-d$ in this case, we

can apply the Chapman-Kolmogorov equations to obtain

$$\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}p_{ij}(s,t,u,d)-p_{ij}(s,t,u,d-h)\\
&=\lim_{h\to 0}\frac{1}{h}p(s,t,u,(d-h,d])\\
&=\sum_{k\in E}\lim_{h\to 0}\frac{1}{h}\int_0^{u+t-d-s}p_{kj}(t-d,t,v,(d-h,d])p_{ik}(s,t-d,u,\,\mathrm{d}v), \qquad \text{(A.60)}
\end{aligned}$$

insofar as the limits exist. We wish to argue that for each k in the sum above, the limit exists. To this end, we first consider the case $k=j$. Note that if $U_{t-d}=v$ for $v>0$ and there are no jumps on $(t-d,t]$, then $U_t>d$. Therefore, we obtain

$$\begin{aligned}
&p_{jj}(t-d,t,v,(d-h,d])\\
&=P(Z_t=j,d-h<U_t\le d|Z_{t-d}=j,U_{t-d}=v)\\
&=P(N_t-N_{t-d}\ge 2,Z_t=j,d-h<U_t\le d|Z_{t-d}=j,U_{t-d}=v). \qquad \text{(A.61)}
\end{aligned}$$

Here, Lemma 3.2 yields

$$\begin{aligned}
&\limsup_{h\to 0}\frac{1}{h}P(N_t-N_{t-d}\ge 2,Z_t=j,d-h<U_t\le d|Z_{t-d}=j,U_{t-d}=v)\\
&\le\limsup_{h\to 0}\frac{1}{h}P(N_t-N_{t-d}\ge 2|Z_{t-d}=j,U_{t-d}=v)=0. \qquad \text{(A.62)}
\end{aligned}$$

Therefore, applying Lemma 3.1 and the dominated convergence theorem, we obtain

$$\lim_{h\to 0}\frac{1}{h}\int_0^{u+t-d-s}p_{jj}(t-d,t,v,(d-h,d])p_{ij}(s,t-d,u,\,\mathrm{d}v)=0. \qquad \text{(A.63)}$$

Next, consider the case $k\neq j$. In this case, Lemma A.7 yields that $p_{kj}(t-d,t,v,\cdot)$ is concentrated on $[0,d]$. Also note that when $U_t>d-h$, no jumps are made in the time interval $(t-d+h,t]$. Therefore, any jumps made by the process in the interval $(t-d,t]$ must in this case be made in the interval $(t-d,t-d+h]$, yielding

$$\begin{aligned}
&p_{kj}(t-d,t,v,(d-h,d])\\
&=P(Z_t=j,d-h<U_t|Z_{t-d}=k,U_{t-d}=v)\\
&=P(N_{t-d+h}-N_{t-d}=1,Z_t=j,d-h<U_t|Z_{t-d}=k,U_{t-d}=v)\\
&+P(N_{t-d+h}-N_{t-d}\ge 2,Z_t=j,d-h<U_t|Z_{t-d}=k,U_{t-d}=v).
\end{aligned}$$

As in the previous case, Lemma 3.2 yields

$$\begin{aligned}
&\limsup_{h\to 0}\frac{1}{h}P(N_{t-d+h}-N_{t-d}\ge 2,Z_t=j,d-h<U_t|Z_{t-d}=k,U_{t-d}=v)\\
&\le\limsup_{h\to 0}\frac{1}{h}P(N_{t-d+h}-N_{t-d}\ge 2|Z_{t-d}=k,U_{t-d}=v)=0, \qquad \text{(A.64)}
\end{aligned}$$

with the convergence, as before, being bounded above by a constant according to Lemma 3.1. Next, note that on $N_{t-d}=n$, having $N_{t-d+h}-N_{t-d}=1$ and $d-h<U_t$ is equal to having $t-d<T_{n+1}\le t-d+h$ and $T_{n+2}>t$. Let

$g(n) = P(N_{t-d} = n | Z_{t-d} = k, U_{t-d} = v)$ and $A = (t-d, t-d+h]$, we therefore obtain

$$
\begin{aligned}
&P(N_{t-d+h} - N_{t-d} = 1, Z_t = j, d-h < U_t | Z_{t-d} = k, U_{t-d} = v) \\
&= \sum_{n=0}^{\infty} P(T_{n+1} \in A, T_{n+2} > t, Z_t = j | Z_{t-d} = k, U_{t-d} = v, N_{t-d} = n) g(n). \qquad \text{(A.65)}
\end{aligned}
$$

Here, by Lemma A.4 and Lemma A.3, the conditional distribution of $(T_{n+1}, Z_{T_{n+1}})$ given $Z_{t-d} = k$ $U_{t-d} = v$ and $N_{t-d} = n$ has density $(l, s) \mapsto r_{kl}(t-d, v, s)$ for $l \neq k$ and $s \geq t-d$, and by Lemma A.2, the survival function of the conditional distribution of T_{n+2} given $T_{n+1} = s$ and $Z_{T_{n+1}} = l$ is $w \mapsto R_l(s, 0, w)$. Here, we use the notation outlined in (A.57) and (A.58). Thus, we obtain

$$
\begin{aligned}
&P(T_{n+1} \in A, T_{n+2} > t, Z_t = j | Z_{t-d} = k, U_{t-d} = v, N_{t-d} = n) \\
&= \int_{t-d}^{t-d+h} r_{kj}(t-d, v, s) R_j(s, 0, t) \, \mathrm{d}s. \qquad \text{(A.66)}
\end{aligned}
$$

As this is independent of n, insertion in (A.65) yields

$$
\begin{aligned}
&P(N_{t-d+h} - N_{t-d} = 1, Z_t = j, d-h < U_t | Z_{t-d} = k, U_{t-d} = v) \\
&= \int_{t-d}^{t-d+h} R_j(s, 0, t) r_{kj}(t-d, v, s) \, \mathrm{d}s. \qquad \text{(A.67)}
\end{aligned}
$$

By right-continuity of the intensities, this yields

$$
\begin{aligned}
&\lim_{h \to 0} \frac{1}{h} P(N_{t-d+h} - N_{t-d} = 1, Z_t = j, d-h < U_t | Z_{t-d} = k, U_{t-d} = v) \\
&= r_{kj}(t-d, v, t-d) R_j(t-d, 0, t) \\
&= q_{kj}(t-d, v) \exp\left(- \int_{t-d}^{t} q_j(r, r-(t-d)) \, \mathrm{d}r \right). \qquad \text{(A.68)}
\end{aligned}
$$

For brevity, we let $a_{kj}(t, d, u, v)$ denote the right-hand side of (A.68). Now using (A.63) and (A.68) in (A.60), we obtain

$$
\begin{aligned}
&\lim_{h \to 0} \frac{1}{h} p_{ij}(s, t, u, d) - p_{ij}(s, t, u, d-h) \\
&= \sum_{k \neq j} \lim_{h \to 0} \frac{1}{h} \int_0^{u+t-d-s} p_{kj}(t-d, t, v, (d-h, d]) p_{ik}(s, t-d, u, \mathrm{d}v) \\
&= \sum_{k \neq j} \int_0^{u+t-d-s} a_{kj}(t, d, u, v) p_{ik}(s, t-d, u, \mathrm{d}v), \qquad \text{(A.69)}
\end{aligned}
$$

yielding (4.4). Note that in order to move the limit under the integral in (A.69), we applied the dominated convergence theorem, making use of that the integral in (A.67) is bounded from above on $[t-d, t-d+\varepsilon]$ for some $\varepsilon > 0$. This shows that $p_{ij}(s, t, u, d)$ is differentiable from the left in d for $d > 0$ when $i \neq j$. The diagonal case follows immediately from this by writing $p_{ii}(s, t, u, d)$ as one minus the sum of $p_{ij}(s, t, u, d)$ for $j \neq i$. $\square$

Proof of Theorem 4.3. Let $s, u \geq 0$ and let $t \geq s$ and $d > 0$. Applying the Chapman-Kolmogorov equations of Lemma A.8, we have

$$\begin{aligned}
&\frac{\partial p_{ij}}{\partial t}(s,t,u,d) \\
&= \lim_{h\to 0}\frac{1}{h}(p_{ij}(s,t+h,u,d) - p_{ij}(s,t,u,d)) \\
&= \lim_{h\to 0}\frac{1}{h}\sum_{k\in E}\int_0^{u+t-s} p_{kj}(t,t+h,v,d)p_{ik}(s,t,u,\,\mathrm{d}v) - \frac{1}{h}p_{ij}(s,t,u,d), \qquad \text{(A.70)}
\end{aligned}$$

insofar as the limits exist, our objective is to argue that this is the case. Now, for $k \neq j$, Lemma A.7 shows that the measure $p_{kj}(t,t+h,v,\cdot)$ is concentrated on $[0,h]$. Therefore, whenever $d \geq h$, we have $p_{kj}(t,t+h,v,d) = p_{kj}(t,t+h,v)$. Applying Lemma 3.4, Lemma 3.3 and the dominated convergence theorem, we then obtain

$$\begin{aligned}
&\lim_{h\to 0}\int_0^{u+t-s}\frac{1}{h}p_{kj}(t,t+h,v,d)p_{ik}(s,t,u,\,\mathrm{d}v) \\
&= \int_0^{u+t-s}\lim_{h\to 0}\frac{1}{h}p_{kj}(t,t+h,v)p_{ik}(s,t,u,\,\mathrm{d}v) \\
&= \int_0^{u+t-s} q_{kj}(t,v)p_{ik}(s,t,u,\,\mathrm{d}v). \qquad \text{(A.71)}
\end{aligned}$$

Now consider the case $k = j$. We wish to evaluate the limit

$$\lim_{h\to 0}\int_0^{u+t-s}\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{1}{h}p_{ij}(s,t,u,d). \qquad \text{(A.72)}$$

We note that for $v \geq d$, Lemma 4.1 yields

$$\limsup_{h\to 0}\frac{1}{h}p_{jj}(t,t+h,v,d) \leq \limsup_{h\to 0}\frac{1}{h}p_{jj}(t,t+h,v,(v+h)-) = 0, \qquad \text{(A.73)}$$

so that $p_{jj}(t,t+h,v,d)/h$ tends to zero as h tends to zero from above. The bound in Lemma 4.1 and the dominated convergence theorem then allows us to conclude that in the case $d < u+t-s$, we have

$$\lim_{h\to 0}\int_d^{u+t-s}\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v) = 0. \qquad \text{(A.74)}$$

Therefore, we obtain that if $d < u+t-s$, then the limit (A.72) is equal to the limit

$$\lim_{h\to 0}\int_0^d\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{1}{h}p_{ij}(s,t,u,d), \qquad \text{(A.75)}$$

provided that this exists. In the case $d \geq u+t-s$, we we note that according to Lemma A.7, $p_{ij}(s,t,u,\cdot)$ is concentrated on $[0,u+t-s]$, so the limits (A.72) and (A.75) are also equal in this case, whenever they exist.

We proceed with evaluating (A.75). Note that $p_{jj}(t,t+h,v,\cdot)$ is concentrated on $[0,v+h]$ by Lemma A.7. Therefore, $p_{jj}(t,t+h,v,d) = p_{jj}(t,t+h,v)$ for

$0 \le v \le d-h$, for $d-h < v \le d$, we have

$$\begin{aligned} p_{jj}(t,t+h,v,d) &= p_{jj}(t,t+h,v,v+h) - p_{jj}(t,t+h,v,(d,v+h]) \\ &= p_{jj}(t,t+h,v) - p_{jj}(t,t+h,v,(d,v+h]). \end{aligned} \tag{A.76}$$

As a consequence, we find that

$$\begin{aligned} &\int_0^d \frac{1}{h} p_{jj}(t,t+h,v,d) p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{1}{h} p_{ij}(s,t,u,d) \\ &= \int_0^d \frac{1}{h} p_{jj}(t,t+h,v) p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{1}{h} p_{ij}(s,t,u,d) \\ &- \frac{1}{h} \int_{d-h}^d p_{jj}(t,t+h,v,(d,v+h]) p_{ij}(s,t,u,\,\mathrm{d}v), \end{aligned} \tag{A.77}$$

where the final integral is over $(d-h,d]$. Taking the limit of the first two terms in (A.77), we may apply the dominated convergence theorem as in the off-diagonal case and obtain

$$\begin{aligned} &\lim_{h\to 0} \int_0^d \frac{1}{h} p_{jj}(t,t+h,v) p_{ij}(s,t,u,\,\mathrm{d}v) - \frac{1}{h} p_{ij}(s,t,u,d) \\ &= \lim_{h\to 0} \int_0^d \frac{1}{h} (p_{jj}(t,t+h,v) - 1) p_{ij}(s,t,u,\,\mathrm{d}v) \\ &= \int_0^d q_{jj}(t,v) p_{ij}(s,t,u,\,\mathrm{d}v). \end{aligned} \tag{A.78}$$

As regards the term being subtracted in (A.77), we can write

$$\begin{aligned} &\frac{1}{h} \int_{d-h}^d p_{jj}(t,t+h,v,(d,v+h]) p_{ij}(s,t,u,\,\mathrm{d}v) \\ &= \frac{1}{h} (p_{ij}(s,t,u,d) - p_{ij}(s,t,u,d-h)) \\ &+ \frac{1}{h} \int_{d-h}^d (p_{jj}(t,t+h,v,(d,v+h]) - 1) p_{ij}(s,t,u,\,\mathrm{d}v). \end{aligned} \tag{A.79}$$

For the first term in (A.79), we may use Lemma 4.2 to obtain that the limit as h tends to zero exists and is given by

$$\lim_{h\to} \frac{1}{h} (p_{ij}(s,t,u,d) - p_{ij}(s,t,u,d-h)) = \frac{\partial p_{ij}}{\partial d}(s,t,u,d), \tag{A.80}$$

where the right-hand side denotes the left derivative. As for the second term in (A.79), we have

$$\begin{aligned}
&\left|\frac{1}{h}\int_{d-h}^{d}(p_{jj}(t,t+h,v,(d,v+h])-1)p_{ij}(s,t,u,\,\mathrm{d}v)\right| \\
&= \left|\int_{d-h}^{d}\frac{1}{h}(p_{jj}(t,t+h,v)-1)+\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v)\right| \\
&\leq \left|\int_{d-h}^{d}\frac{1}{h}(p_{jj}(t,t+h,v)-1)p_{ij}(s,t,u,\,\mathrm{d}v)\right| \\
&+ \left|\int_{d-h}^{d}\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v)\right|. \qquad \text{(A.81)}
\end{aligned}$$

Here, the integrand in the first term of the final estimate converges to $q_{jj}(t,v)$, and by Theorem 3.3, we also have for $d-h<v\leq d$ that

$$\frac{1}{h}(p_{jj}(t,t+h,v)-1)\leq \sup_{d-h<v\leq d} C(t,u), \qquad \text{(A.82)}$$

the latter being finite. Therefore, the dominated convergence theorem yields that the first term converges to zero. As regards the second integrand, we note that $d-h<v\leq d$ implies $d<v+h$, so Lemma 4.1 yields that the integrand in this case tends to zero, dominated by a bound measurable in t and u and bounded on compacts. Therefore, the dominated convergence theorem also allows us to conclude that the second integral also tends to zero. All in all, this yields

$$\lim_{h\to 0}\left|\int_{d-h}^{d}\frac{1}{h}(p_{jj}(t,t+h,v,(d,v+h])-1)p_{ij}(s,t,u,\,\mathrm{d}v)\right|=0. \qquad \text{(A.83)}$$

Combining (A.83) with our previous conclusions and simplifications, we finally conclude

$$\begin{aligned}
&\lim_{h\to 0}\int_{0}^{u+t-s}\frac{1}{h}p_{jj}(t,t+h,v,d)p_{ij}(s,t,u,\,\mathrm{d}v)-\frac{1}{h}p_{ij}(s,t,u,d) \\
&= \int_{0}^{d} q_{jj}(t,v)p_{ij}(s,t,u,\,\mathrm{d}v)-\frac{\partial p_{ij}}{\partial d}(s,t,u,d), \qquad \text{(A.84)}
\end{aligned}$$

and from (A.71) and (A.84), the theorem follows. □

References

1. K. B. Athreya, D. McDonald, and P. Ney, *Limit Theorems for Semi-Markov Processes and Renewal Theory for Markov Chains*, Ann. Prob. **6** (1978), no. 5, 788–797.
2. K. Buchardt, T. Møller, and K. B. Schmidt, *Cash flows and policyholder behaviour in the semi-Markov life insurance setup*, Scand. Act. Journ. (2014), 1–29.
3. M. Christiansen, *Multistate models in health insurance*, Adv. Stat. Anal. (2012), 155–186.
4. E. Cinlar, *Markov renewal theory: A survey*, Management Science (1975), no. 21, 727–752.

5. G. D'amico, F. Petroni, and F. Prattico, *Semi-Markov Models in High Frequency Finance: A Review*, Stock Markets: Emergence, Macroeconomic Factors and Recent Developments, Nova Science Publishers, 2013, pp. 175–192.
6. ______, *Wind speed modeled as an indexed semi-Markov process*, Environmetrics **6** (2013), 367–376.
7. W. Feller, *On semi-Markov processes*, Proc. Natl. Acad. Sci. (1964), 653–659.
8. Willy Feller, *On the Integro-Differential Equations of Purely Discontinuous Markoff Processes*, Trans. Amer. Math. Soc. **48** (1940), no. 3, 488–515.
9. M. Helwich, *Durational effects and non-smooth semi-Markov models in life insurance*, Universität Rostock, 2007, Ph. D. thesis.
10. J. M. Hoem, *Inhomogeneous Semi-Markov Processes, Select Actuarial Tables and Duration-Dependence in Demography*, Working paper series (1972), 1–36, `http://www.ssb.no/a/histstat/ano/ano_io72_11.pdf`.
11. J. J. Hunter, *On the Moments of Markov Renewal Processes*, Working paper series (1968), 1–33, `http://www.jstor.org/stable/1426217`.
12. Martin Jacobsen, *Point process theory and applications*, Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 2006.
13. Jean Jacod and Albert N. Shiryaev, *Limit theorems for stochastic processes*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer-Verlag, Berlin, 2003.
14. J. Janssen and R. De Dominicis, *Finite non-homogeneous semi-Markov processes*, Insur. Math. Econ. **3** (1984), 157–165.
15. J. Janssen and R. Manca, *Applied Semi-Markov Processes*, Springer-Verlag, London, 2006.
16. ______, *Semi-Markov Risk Models for Finance, Insurance and Reliability*, Springer-Verlag, New York, 2007.
17. X. et al Jianhong, *Spatial and temporal modeling of tourist movements using semi-Markov processes*, Tourism Management **32** (2011), no. 4, 844–851.
18. P. Lévy, *Processus semi-Markoviens*, Proc. Int. Congress. Math. (Amsterdam), vol. 3, 1954, pp. 416–426.
19. A. et al Lucas, *Nonparametric Estimation for Non-Homogeneous Semi-Markov Processes: An Application to Credit Risk*, Preprint (2006), 1–41, Tinbergen Institute Discussion Paper No. 06-024/2. `http://ssrn.com/abstract=892782`.
20. U. U. Mueller, P. E. Greenwood, and W. Wefelmeyer, *Efficient estimation for semiparametric semi-Markov processes*, Comm. Stat. Theo. Meth. (2004), 1–15.
21. J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original.
22. R. Pyke, *Markov renewal processes: Definitions and preliminary properties*, Ann. Math. Stat. **32** (1961), no. 4, 1231–1242.
23. R. Pyke and R. Schaufele, *Limit Theorems for Markov Renewal Processes*, Ann. Math. Statist. **35** (1964), no. 4, 1746–1764.
24. W. L. Smith, *Regenerative stochastic processes*, Proc. Roy. Soc. Lon. A (1955), 6–31.
25. G. Vassiliou and A. Papadopoulou, *Non homogeneous semi-Markov systems and maintainability of state sizes*, Journ. Appl. Prob. **29** (1992), 519–534.
26. O. Vega-Amaya, *A note on the regularity property of semi-Markov proecsses with Borel state space*, Preprint (2002), 1–8, `http://www.mat.uson.mx/depto/publicaciones/reportes/pdf/reporte18.pdf`.
27. J. Yackel, *Limit theorems for semi-Markov processes*, Trans. Amer. Math. **123** (1966), no. 2, 402–424.

Alexander Sokol: Institute of Mathematics, University of Copenhagen, 2100 Copenhagen, Denmark, alexander@math.ku.dk